AKPIK 2: Applications in Particle and Astroparticle Physics

Time: Tuesday 17:00–19:00 Location: ZEU/0118

AKPIK 2.1 Tue 17:00 ZEU/0118

Studies of Machine Learning Inspired Clustering Algorithms for Jets — Amrita Bhattacherjee[1], Debarghya Ghoshdastidar[1], Stefan Kluth[2], and •Siddha Hill[2] — [1]Technical University of Munich — [2]Max Planck Institute for Physics, Munich

We study several machine learning inspired clustering algorithms to cluster the particles of hadronic final states in high energy e+e- and pp collisions into jets. We compare their performance against well known algorithms such as JADE or Anti-Kt. Performance indicators are physically motivated and study properties such as energy and angle differences of jets transitioning from parton to hadron level. In addition we also investigate the stability against pileup.

AKPIK 2.2 Tue 17:15 ZEU/0118

Providing GPU resources in a HEP analysis environment — Johannes Erdmann[1], Benjamin Fischer[1], Thomas Kress[2], Dennis Noll[1], Andreas Nowack[2], and •Roman Suveyzdis[1] — [1]III. Physikalisches Institut A, RWTH Aachen University — [2]III. Physikalisches Institut B, RWTH Aachen University

Graphics Processing Units (GPUs) have become a key computing resource for advanced physics analyses for example for training and evaluating machine learning models. The local research group's computing cluster is still often limited in GPU resources. Therefore, an efficient and fair use is crucial. HTCondor, a powerful batch job software system that is often used in the HEP physics community, offers the possibility to manage GPU resources. The basic installation of HTCondor only distributes entire GPUs on a per job basis, which leaves some potential resources unused. In this talk, we will present how HTCondor can be configured to cope with the users' need to use GPUs both interactively and in batch job mode. We will report on the first experiences with our setup.

AKPIK 2.3 Tue 17:30 ZEU/0118

Fast Columnar Physics Analyses of Terabyte-Scale LHC Data on a Cache-Aware Dask Cluster — Svenja Diekmann, Niclas Eich, Martin Erdmann, Peter Fackeldey, •Benjamin Fischer, Dennis Noll, and Yannik Rath — III. Physikalisches Institut A, RWTH Aachen University

The development of an LHC physics analysis involves numerous investigations that require the repeated processing of terabytes of data. Thus, a rapid completion of each of these analysis cycles is central to mastering the science project.

We present a solution to efficiently handle and accelerate physics analyses on small-size institute clusters. Our solution uses three key concepts: Vectorized processing of collision events, the "MapReduce" paradigm for scaling out on computing clusters, and efficiently utilized SSD caching to reduce latencies in IO operations. This work focuses on the latter key concept, its underlying mechanism, and its implementation.

Using simulations from a Higgs pair production physics analysis as an example, we achieve an improvement factor of 6.3 in the runtime for reading all input data after one cycle and even an overall speedup of a factor of 14.9 after 10 cycles, reducing the runtime from hours to minutes.

AKPIK 2.4 Tue 17:45 ZEU/0118

ProGamer: PROgressively Growing Adversarial Modified (transformer-)Encoder Refinement — •Benno Käch, Isabell Melzer-Pellmann, and Dirk Krücker — Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany

Machine learning-based data generation has become a major research topic in particle physics due to the computational challenges posed by current Monte Carlo simulation approaches for future colliders, which will have significantly higher luminosity. The generation of collider data is similar to point cloud generation, but it is more difficult because of the complex correlations that need to be accurately modeled between the points. A refinement model consisting of normalising flows and transformer encoders is presented. The normalising flow is 3-dimensional, meaning that the generated particle cloud consists of independent and identically distributed objects. This output is then refined by a transformer encoder, which is adversarially trained against another transformer encoder discriminator/critic. As the model is able to produce an arbitrary number of particles, a progressively growing point cloud can be produced.

AKPIK 2.5 Tue 18:00 ZEU/0118

Machine Learning based defect detection for large-scale electrodes — •Sebastian Vetter — Karlsruhe Institute of Technology, Institute for Astroparticle Physics

Like every piece of hardware produced in an industrial setting, detectors in physics experiments are subject to material defects, introduced during the production or handling of individual components. This can greatly influence detector behavior and lead to unexpected experimental results, depending on the affected part and the extent of the defect. Detection and quantification of such defects is therefore an important step in constructing a successful experiment.

It is still quite common for defect inspection to be done by eye. However, recent developments in computer-based inspection methods provide the opportunity to relieve humans from this tedious task, to remove the susceptibility of human error from the inspection step, and to objectively quantify the extent of detected defects.

In this talk, I present the defect inspection of a large-scale electrode mesh, as used for example in liquid noble gas Dark Matter experiments. This inspection was carried out first by hand and then compared to various Machine Learning approaches, ranging from simple decision trees to variational autoencoders.

AKPIK 2.6 Tue 18:15 ZEU/0118

Interpolation of Instrument Response Functions for the Cherenkov Telescope Array — •Rune Michael Dominik and Maximilian Linhoff for the CTA-Collaboration — Astroparticle Physics WG Elsässer, TU Dortmund University, Germany

The Cherenkov Telescope Array (CTA) will be the next generation ground-based very-high-energy gamma-ray observatory, utilizing tens of Imaging Atmospheric Cherenkov Telescopes at two sites once its construction and commissioning is finished. Like its predecessors, CTA relies on Instrument Response Functions (IRFs) to relate the observed and reconstructed properties to the original properties of primary particles. IRFs are needed for the proper reconstruction of spectral and spatial information and are thus among the data-products issued to the observatory's users. They are derived from Monte Carlo simulations and depend on observation conditions like the telescope pointing direction or the atmospheric transparency. Producing a complete set of IRFs from simulations for every observation taken is a time consuming task and not feasible when releasing data-products on short timescales. Consequently, interpolation techniques on simulated IRFs are investigated to quickly estimate IRFs for specific observation conditions. However, as some of an IRFs constituents are given as probability distributions, specialized methods are needed. This talk summarizes and compares the feasibility of multiple approaches to interpolate IRF components. First results are shown and open challenges are discussed.

AKPIK 2.7 Tue 18:30 ZEU/0118

Estimation of prediction uncertainties for data from Imaging Atmospheric Cherenkov Telescopes — •Cyrus Pan Walther and Maximilian Linhoff — Technische Universität Dortmund, Germany

One main step in the low-level analysis of astroparticle physics data is the reconstruction of the properties of primary particles that induced extensive air showers.

Various methods are applied in different experiments and software packages. In general, these are multi-output and combined regression and classification tasks. The estimation of prediction uncertainties is of crucial importance for the later scientific exploitation of these events. However, most methods do not in themselves provide reliable uncertainty estimates. In this contribution, we want to apply a method that has been used successfully in a Deep Learning reconstruction for the IceCube experiment to data from Imaging Atmospheric Cherenkov Telescopes used for gamma-ray astronomy.

AKPIK 2.8 Tue 18:45 ZEU/0118

Testing Nested Machine Learning Models for the Cherenkov Telescope Array — •Lukas Beiske and Rune M. Dominik for the

CTA-Collaboration — Astroparticle Physics, WG Rhode/Elsässer, TU Dortmund University, D-44227 Dortmund, Germany

The Cherenkov Telescope Array (CTA) will be the next-generation ground-based very-high-energy gamma-ray observatory covering an energy range from 20 GeV up to 300 TeV. It will operate tens of Imaging Atmospheric Cherenkov Telescopes (IACTs) on the Canary Island of La Palma (CTA North) and at the Paranal Observatory in Chile (CTA South) once construction and commissioning are finished.

Machine Learning techniques are currently being used to analyze data from IACTs. The tools are used to reconstruct the three main properties of the primary particle: its particle type, energy, and origin. A common approach is to train models on parameters extracted from the shower images observed by the telescopes which in turn give one prediction per telescope image. For events triggering multiple telescopes, these individual predictions can be averaged to obtain a single primary particle prediction for every shower event. However, it is possible to improve these averaged predictions by training a second set of machine learning models using all information available about the shower as seen by the whole telescope array. This talk will show the current results of testing such nested models for CTA.